



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

TESTS AND MEASUREMENTS IN MODERN LANGUAGE WORK

By C. H. HANDSCHIN

WE HAVE reached a point in our modern language teaching where we feel the need of standard objective scales by which to measure the product of our teaching. In some lines, especially in elementary school teaching, the work in measurements has been carried to the point of considerable perfection, and we recognize that it means a decided advance in our educational work. In modern languages we are just making a start in this direction.

Now before we can measure, we must have something to measure. To get something to measure is therefore the most important thing, and good, painstaking teaching is still the all important factor. But we know also that measuring results exactly gives us a check on our work, shows us its weak and strong points and stimulates both student and teacher. Scientifically inclined teachers desire exact measurements. In language, the phenomena treated are more intangible than in many other subjects. The results seem discouragingly intangible often. We desire and we need to know what we have accomplished. We can learn this best by thoroughly objective scientific standards of measurements.

I cannot go into a discussion of the laboratory and classroom experiments which have been conducted, chiefly in Europe, with a view to getting light on the language-learning process. Suffice it here to give a slight resumé of the principles which we may consider as having been sufficiently established.

Resumé

The following aims and principles may be considered as sufficiently established to warrant their use in pedagogical practice:

1. The fourfold process of learning language, i.e., hearing, speaking, seeing (reading) and writing—the first of which must be aural, the second oral.

2. Skill in motor control (speaking and writing) and in visual perception, and discipline in good habits of study. (Wundt,¹ Eggert,² *et al.*)

3. There must be conscious memorial activity; matter which is to be retained must be repeated, i.e., presented repeatedly to consciousness; rhythmical form, or recurrence, favors retention; attention is an important factor in memorial activity, and attention depends on interest; the feelings (of pleasure or pain) also play a great part in memorial activity and the feelings manifest themselves in the form of interest. (Meumann³ *et al.*)

4. The memory for objects and movements is greater than for verbal impressions. (Peterson,⁴ Kirkpatrick,⁵ Calkins,⁶ Pohlmann.⁷)

5. The ability to recall the vernacular word on presentation of the foreign word (translation from the foreign language) is much greater (two or three times as great) as the ability to recall the foreign word. (Schuyten.⁸) Both foreign word-native word and native word-foreign word learning are superior to teaching foreign words by means of pictures in point of easiest and surest

¹ Psychologische Physiologie. Vol. 1, 170 ff., 238 ff.

² Eggert, Der Psychologische Zusammenhang in der Didaktik des neu-sprachlichen Unterrichts. p. 43 ff.

³ Meumann: The Psychology of Learning, Appleton, 1913, p. 15 ff. and *passim*.

⁴ Peterson: Recall of words, objects and movements, Psychological Review. Monograph Supplement 4, 207-33 (1903).

⁵ E. A. Kirkpatrick: An experimental Study of Memory. Psychological Review 1, 602-609 (1894).

⁶ M. W. Calkins: A Study of Immediate and Delayed Recall of the Concrete and of the Verbal. Psychological Review 5, 451-462 (1898).

⁷ A. Pohlmann: Experimentelle Beiträge zur Lehre von Gedächtnis. 1906, p. 71 ff. See also *ibid.*, 145 ff.

⁸ M. C. Schuyten: Experimentelles zum Studium der gebräuchlichsten Methoden des fremdsprachlichen Unterrichts. Zeitschrift für experimentelle Pädagogik, 3, 199-211 (1906).

retention, of fatigue, most ready reproduction, and of dependence upon form of learning. (Netschajeff.⁹)

6. The object-foreign word method of learning is superior to the foreign word-native word method, and this is superior to the native word-foreign word method in point of immediate as well as permanent retention. (Braunshausen.¹⁰)

7. Learning words in sentences is easier for immediate or deferred recall than learning isolated words. (Libby,¹¹ Grinstead,¹² Binet and Henri.¹³) It depends upon the nature of the test (uses to which the knowledge is put) as to which mode of presentation is best employed (Schlüter.¹⁴)

8. "The study of foreign languages materially increased the student's knowledge of English grammar, but only slightly increases his ability to use English correctly." (Starch.¹⁵)

9. "Training in foreign language seems to have produced a distinct effect in greater fluency of words in writing and in more rapid perception of words in reading." (Starch.¹⁶)¹⁷

⁹ Psychologische Betrachtungen zur Frage über den fremdlandischen Sprachunterricht. Pädagogisch psychologische Studien, 9, Nos. 1 and 2 (1908).

¹⁰ Les méthodes d'enseignement des langues étrangères. Revue Psychologique (1910), Vol. 3.

¹¹ W. Libby: An Experiment in Learning a Foreign Language, Pedagogical Seminar, 17, 81-96 (1910).

¹² W. J. Grinstead: An Experiment in Learning Foreign Words. Journal of Educational Psychology, 6, 242-245 (1915).

¹³ A. Binet et V. Henri: La Mémoire des Mots. Année Psychologique, 1, 1-23 (1894).

¹⁴ L. Schlüter: Experimentelle Beiträge Zur Prüfung der Anschauungs- und der Übersetzungsmethode bei der Einführung in einen fremdsprachlichen Wortschatz. Zeitschrift für Psychologie, 68, Nos. 1 and 2 (1911).

¹⁵ D. Starch: Some Experimental Data on the Value of Studying Foreign Languages. School Review, 23, 697-704 (1915).

¹⁶ *Idem.*, The Value of Studying Foreign Languages, *ibid.*, 25, 243-249 (1917).

¹⁷ High probability is established by Miss Clarahan for the following principle: The reading method is superior to the grammar-translation method for assimilating reading texts as well as for assimilating grammatical knowledge. Cf., M. Clarahan. An Experimental Study of Methods of Teaching High School German. Bulletin of the University of Missouri. Educational Series, Vol. I, No. 6, 1913.

As to attempts to construct scales for measuring the results of foreign language teaching, there are only two to be recorded; one for Latin by Professor P. Hanus,¹⁸ which tests vocabulary, translation and grammar: words and sentences taken from Caesar and Cicero and grammar based on these sentences. Very interesting. Hard to score. Not standardized.

The second attempt is that of Professor Starch, who has set up two tentative tests applicable to Latin, French, or German published in his "Educational Measurements," MacMillan, 1917. The one is a vocabulary test. Two lists of 100 words each, got for French by choosing the first word on every sixth or seventh page of Spiers' and Surenné's large French-English Dictionary; for German by taking the first word on every twenty-third page of the large Müret-Sanders German-English Dictionary. These lists are accompanied by the list of the English equivalents of the foreign words. The student's task is to match the two.

That this test can be of little value to modern language teachers or students should it ever be standardized is quite plain, since it tests a promiscuous vocabulary, which we do not seek to teach and since the student wastes valuable time in doing the work. The test is the same *mutatis mutandis* for French, German and Latin.

The second test is a reading test, so named, but is in reality a translation test, and consists of thirty sentences graded in difficulty, according to the *judgement* of the inventor of the test. In other words, it has not been standardized.

These strictures on the work of a fellow-worker sound harsh but they are made for the following further reasons:

As to Test 1: A knowledge of lists of isolated words is a poor criterion by which to judge a well trained modern language student who has learned his vocabulary in context. As to the Reading Test: It calls for an ability which we do not consider a major aim. The translation according to his scheme of scoring is either entirely right or entirely wrong. But this principle does not hold for scaling translation as it does in scoring the answers to problems, or to a question purposely so worded that the answer can be only right or wrong.

¹⁸ Progress in Learning Latin (See School Review, 24, 342-51 (1916).

Let us inquire next into the principles which should obtain in making scales. The test must be: (1) objective, i.e., must not depend on any peculiar training of the student but should measure what is generally considered desirable in student-training and what is generally taught. Further, the interpretation of the students' work must admit of no subjective or individual opinion, i.e., the form in which the material is called for and the method of scoring must be such that half a dozen or any number of teachers administering the test separately will each get the same reaction and the same score. (2) It must test in a manner analogous to the one usually employed by teachers, and not in a way in which classes have never reacted before, which would be obviously unjust and unwise. (3) It must be of such a nature as to be valuable to both student and teacher from the start, i.e., even before it has been standardized. (4) It must be comprehensive enough to include all from the poorest to the best students. (5) The units of the scale must not be too large nor too small, especially where the answer is by its nature entirely right or entirely wrong. (6) It must be tested out in several thousand cases, at least. (7) It must be simple, so that teacher and student will not waste time attempting to learn what is wanted. (8) And perhaps most important, if many teachers are to be induced to use it and use it over and over year after year, the method of administering it and, especially, of scoring must be simple, and require very little time.

Coming to the tests which I wish particularly to note, the following Reading Test B is given by way of illustration.

Silent Reading Test B

FOR FIRST OR SECOND-YEAR FRENCH IN A FOUR-YEAR HIGH SCHOOL

Name..... Date..... Score.....

DIRECTIONS. Sign your name and fill in the date above. Read the text found on the other side of this sheet as rapidly as possible, but be sure you get the meaning, and read it with the intention of answering in English or French the questions appended. When I give the word "Start," turn over this sheet and study until I give the word "Stop," this will give you one minute for study. Then draw a circle around the last word you read, turn the sheet back over, and begin at once to answer in

English, or French, the ten questions found on this sheet. You will be allowed five minutes to answer the questions.

QUESTIONS ON THE TEXT

(Answer in English or French as you prefer)

1. Of what two characters does the story treat?
2. What did they finally agree to do?
3. Who had signed the preliminaries of the peace treaty?
4. What was the most important article of the treaty?
5. That he might not eat the young owls, what was it necessary for the eagle to know?
6. Where did the eagle finally see the young owls?
7. What did he say to himself when he saw them?
8. What did he set about doing then?
9. In what respect did the young owls agree with the old owl's description of them?
10. What does this weakness on the part of parents toward their children often do?

11 L'aigle et le hibou, après avoir fait longtemps la guerre con-
 20 vinrent d'une paix; les articles préliminaires avaient été
 29 préalablement signés par des ambassadeurs: l'article le plus
 40 essentiel était que le premier ne mangerait pas les petits de
 48 l'autre. Les connaissez-vous? demanda le hibou. Non, ré-
 60 pondit l'aigle. Tant pis. Peignez-les-moi ou montrez-les-moi.
 71 Foi d'honnête aigle je n'y toucherai jamais. Mes petits, ré-
 82 pondit l'oiseau nocturne, sont mignons, beaux, bien faits; ils ont
 91 la voix douce et mélodieuse; vous les reconnaîtrez aisément
 104 à ces marques. Très bien, je ne l'oublierai pas. Il arriva un
 117 jour que l'aigle aperçut dans le coin d'un rocher de petits mons-
 128 tres très laids, rechignés, avec un air triste et lugubre. Ces en-
 139 fants, dit-il, n'appartiennent pas à notre ami; mangeons-les;
 154 aussitôt il se mit à en faire un bon repas. L'aigle n'avait pas
 166 tort. Le hibou lui avait fait une fausse peinture de ses petits;
 174 ils n'en avaient pas le moindre trait.

184 Les parents devraient éviter avec soin ce faible envers leurs
 193 enfants, il les rend souvent aveugles sur leurs défauts.

SCORE CARD FOR SILENT READING TEST B, FIRST- OR SECOND-YEAR FRENCH

Teacher..... School.....

First Year French. Date.....

RATE SCORE

Interval	No. of scores	Interval	No. of scores	Interval	No. of scores	Interval	No. of scores	QUALITY SCORE		
								Questions	Number of Rights	Number of Wrongs
91 to 95				91 to 95						
86 to 90				86 to 90						
81 to 85				81 to 85						
76 to 80		Above 160		76 to 80						
71 to 75		156 to 160		71 to 75						
66 to 70		151 to 155		66 to 70		151 to 155				
61 to 65		146 to 150		61 to 65		146 to 150				
56 to 60		141 to 145		56 to 60		141 to 145		10		
51 to 55		136 to 140		51 to 55		136 to 140		9		
46 to 50		131 to 135		46 to 50		131 to 135		8		
41 to 45		126 to 130		41 to 45		126 to 130		7		
36 to 40		121 to 125		36 to 40		121 to 125		6		
31 to 35		116 to 120		31 to 35		116 to 120		5		
26 to 30		111 to 115		26 to 30		111 to 115		4		
21 to 25		106 to 110		21 to 25		106 to 110		3		
16 to 20		101 to 105		16 to 20		101 to 105		2		
Below 16		96 to 100		Below 16		96 to 100		Q1		

NOTE.—If this score card is to be used for *second*-year French strike out the word *First* above and substitute the word *Second*. Columns 5, 6, 7 above are of use only in a combined French and Spanish score-card.

Instructions for Making the Distribution of Pupils' Scores and of Finding the Median Score

Arrange the children's papers for any class in order of the scores, the lowest score on top. To make the distribution called for, count the number of papers whose scores fall within the successive groups listed. For instance, if the lowest score made is 5, the next 17, 26, 30, you will put "1" in the group marked below 16, "2" in the group marked 16 to 20, "3" and "4" in the group marked 26 to 30, etc.

The median score is the score on the middle paper in the pile of papers arranged according to the size of the scores. If there are 25 papers, the median score is the score on the 13th paper,

because there are 12 papers above it in score and 12 papers below it. If there are 26 papers, the median score is half way between the score of the 13th and 14th papers.

Key to the Answers

To be right the answers must express the *exact thought* expressed in the following answers 1. The eagle and the owl. 2. To make peace. 3. The plenipotentiaries. 4. That the eagle should not eat the young of the owl. 5. To know how they looked. 6. In a crevice of a rock. 7. These are not the young of my friend, I shall eat them. (The last thought of the preceding sentence may be omitted.) 9. In no respect. 10. To blind them to their faults.

Method of Scoring

As to *Quality Score* the teacher may for the benefit of his class assign one point for each question answered correctly in order to be able to give each pupil a score. However, for our purposes he will enter merely the total number of times a question was answered "right" and the total number of times it was answered "wrong," in the spaces provided for that purpose on the score card above.

As to the *Rate Score*, the number of words to and including the one around which the pupil has put a circle is his "Rate." Distribute this along with the others as directed above. To facilitate counting the words a student has read, the number of words contained in each line, and all the lines preceding it, is indicated before each line.

Tentative Standards:

Number of pupils: First year French, 67.

Quality score:

Questions answered right: 2 questions.

Rate score: First year French, 52 words.

A similar test is arranged for German and Spanish. A second reading test for French, German, and Spanish is of the problem type. A third type of test, arranged also for the three languages, is the grammar and comprehension test for the first year only. Further tests for advanced classes are being set up.

The standardization and use of tests and measurements is certain to constitute the next important step in modern language

teaching. An appeal is hereby made to teachers to use the tests described above, since it is only through the collective activity of numerous teachers that such tests can be standardized.

A word may be permitted here also concerning the "Test to Discover Types of Learners," published in its first conception in *Modern Language Journal*, Vol. 3, p. 1 ff. The writer has since then continued the work of perfecting this test, i.e., standardizing the material which it employs. Data for this purpose are now being collected in several centers. As soon as it is perfected it will be published. Several outside groups, the writer has been informed, are attempting to perfect this test. In one case which has been brought to his attention, this is being done in a rather crude manner, and without giving due credit to its inventor. The writer considers this test by far the most important one yet devised and the most important for the future of modern language study, since, when perfected, it bids fair to revolutionize our practice in regard to placing students in the classes for which their natural endowments qualify them.

Note to aid the understanding of technical terms: On the score card the word "Interval" indicates that a student's rate score is to be entered after the first interval, "Below 16," if his rate score is below 16, after the second interval, "16 to 20," if his score is anywhere from 16 to 20, etc. Providing for 193 intervals on the card would consume too much space, and the result, statistically, is the same. The caption, "Number of scores," means the number of students whose rate score falls within the intervals stated. The words, "Tentative Standards," mean that the grades or scores there given have been averaged by the numbers of students there given. Since the average of greater numbers of students will no doubt vary from this, the words *tentative standard* (or average) are used.

Note.—The tests for first and second-year French, German, and Spanish have been used for a year. They are now being printed and will soon be available. Address: The World Book Company, Yonkers, New York.

Miami University.